Interdependent preferences

Kapteyn, A.; van der Geer, S.; van de Stadt, H.; Wansbeek, T.

Published in:
Journal of Applied Econometrics

Publication date:
1997

Citation for published version (APA):
Kapteyn, A., van der Geer, S., van de Stadt, H., & Wansbeek, T. (1997). Interdependent preferences: An econometric analysis. *Journal of Applied Econometrics*, *12*(6), 665-686.

JOURNAL OF APPLIED ECONOMETRICS, VOL. 12 665–686 (1997)

INTERDEPENDENT PREFERENCES: AN ECONOMETRIC ANALYSIS

ARIE KAPTEYN,[a*] SARA VAN DE GEER,[b] HUIB VAN DE STADT[c] AND TOM WANSBEEK[d]

[a] *CentER, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: A.kapteyn@kub.nl*
[b] *Department of Mathematics and Computer Science, Leiden University, PO Box 9512, 2300 RA Leiden, The Netherlands*
[c] *Department of Financial Enterprises, Statistics Netherlands, PO Box 959, 2270 AZ Voorburg, The Netherlands*
[d] *Department of Economics, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands*

SUMMARY

The theoretical model of Gaertner (1974) and Pollak (1976) for the interdependence of preferences in the Linear Expenditure System is estimated for a cross-section of households. The interdependence of consumption of different households has implications for the stochastic structure of the model and for the identifiability of its parameters. Both aspects are dealt with. The empirical results indicate a significant role played by the interdependence of preferences. One of its implications is that predictions of the effects of changes in a household's exogenous variables differ according to whether the exogenous variable only changes for this household or for all households jointly.

J. Appl. Econ., **12**, 665–686 (1997)

No. of Figures: 0. No. of Tables: 2. No. of References: 52.

1. INTRODUCTION

In his pioneering study, Duesenberry (1949) gave several kinds of evidence based on aggregate data to indicate the importance of preference interdependence for the explanation of consumer behaviour. At about the same time Leibenstein (1950) extensively discussed various types of interdependencies in consumption behaviour of individuals. Of course, these two authors were not the first to discuss preference interdependence. Leibenstein notes, for example, that the notion of 'conspicuous consumption' can be traced back as far as the works of Horace. Since the time the papers by Duesenberry and Leibenstein were published, some further work has been done on what has been called alternatively variable preferences, endogenous preferences, or interdependent preferences. In Kapteyn *et al.* (1980) we have given a brief review of most of this literature. More recent work includes Frank (1984, 1985, 1989) and Blomquist (1993).

In the economics literature endogenous preferences are usually assumed to arise as a result of habit formation (e.g. Pollak and Wales, 1969; Houthakker and Taylor, 1970; Phlips, 1972, 1974; Manser, 1976; Spinnewijn, 1981; Phlips and Spinnewijn, 1982; Muellbauer, 1988; Pashardes, 1986; Winder and Palm, 1989). Where in the older literature, habit formation was invariably myopic (or 'naive'), in the more recent literature inspired by Spinnewijn (1981) habits are allowed

* Correspondence to: Arie Kapteyn, CentER, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: A.kapteyn@kub.nl

Contract grant sponsor: Netherlands Organization for Scientific Research

CCC 0883–7252/97/060665–22$17·50 *Received 2 June 1993*
 Revised 15 November 1996

to be rational, i.e. consumers may anticipate the effect of their current decisions on their future preferences. The empirical evidence on the extent to which habit formation is myopic or rational is mixed. Pashardes (1986) finds that entirely myopic habits are rejected by his data but fully rational habits are not, whereas Muellbauer (1988) finds habits to be predominantly myopic.

In the context of the Linear Expenditure System (LES), the two forms of habit formation are basically indistinguishable (Spinnewijn, 1981; Phlips and Spinnewijn, 1982). As we will be concerned exclusively with the LES, we can therefore ignore the distinction.

In two rather closely related papers Gaertner (1974) and Pollak (1976) have studied some theoretical implications of the incorporation of preference interdependence in the LES. Darrough, Pollak, and Wales (1983) estimate a Quadratic Expenditure System for three separate time series (one British and two Japanese) of grouped household budget data. They also consider specifications where some parameters depend on lagged consumption. Since the data are grouped according to income-demographic cells, one may interpret the dependence of parameters on lagged consumption (i.e. the consumption of other people in the same cell, one period ago) as representing interdependent preferences. The authors find the specification with lagged consumption included to be empirically superior to static versions of their model.

Alessie and Kapteyn (1991) use a two-wave panel of consumer expenditures to estimate an Almost Ideal Demand System with both habit formation and interdependence. The interdependence is modelled by making current budget shares of a household dependent upon mean budget shares in the reference group of this household. The construction of reference groups requires a careful consideration of underlying assumptions. The assumptions used by Alessie and Kapteyn (1991) are essentially the same as in the present paper. They are spelt out in Section 3 and Appendix A.

The current paper considers interdependence in the LES, exploiting cross-sectional data on consumer expenditures. In contrast to Alessie and Kapteyn (1991), interdependence is modelled by making parameters in the LES dependent upon *current* quantities in the reference group of a household. This feature raises identification issues stemming from the 'reflection problem' formulated by Manski (1993). A major share of the paper is devoted to this.

The paper is organized as follows. Section 2 presents the LES with interdependence incorporated. Section 3 and Appendix A concentrate on the stochastic assumptions required to render the model amenable to estimation on the basis of a single cross-section. In Section 4 we consider issues of identification. Section 5 contains the results of estimating the model for a household expenditure survey in the Netherlands. Section 6 concludes with some qualifications, and points at future research.

2. THE DETERMINISTIC PART OF THE MODEL

Our starting point is the Linear Expenditure System (LES):

$$x_{gn} = b_{gn} + \gamma_g \left(y_n - \sum_{h=1}^{G} b_{hn} \right) \tag{1}$$

where the index n ($n = 1, \ldots, N$) indicates the N consumers (or households; for the purpose of this paper we use the terms 'household', 'family', 'individual', 'consumer' as synonyms) in society; the index g ($g = 1, \ldots, G$) indicates goods; x_{gn} denotes the quantity of good g consumed by individual n; y_n is total expenditures; γ_g, with $\Sigma_g \gamma_g = 1$, and b_{gn} are parameters. System (1) arises from the maximization of the utility function

$$U_n(x_1, \ldots, x_G) = \sum_{g=1}^{G} \gamma_g \log(x_g - b_{gn}) \tag{2}$$

subject to the budget constraint

$$\sum_{g=1}^{G} x_{gn} = y_n \tag{3}$$

with y_n total expenditures of household n. Notice that in (3) the prices of all goods are equal to one. Since we will be dealing with a cross-section where all consumers face the same prices, this does not involve any loss of generality. As a result we will use 'consumption' and 'expenditures' as synonyms.

A maximization problem like this may arise in a rather general setting, like maximization of an intertemporally separable Stone–Geary utility function under a lifetime budget constraint. In such circumstances the maximization problem (2)–(3) arises in the second stage of a two-stage budgeting process. Introduction of certain forms of uncertainty or liquidity constraints does not affect the second stage. See e.g. Blundell and Walker (1986) and Alessie and Kapteyn (1989).

We incorporate interdependence of preferences by expressing the parameters b_{gn} as a function of consumption by others:

$$b_{gn} = b_{g0} + \beta_g \sum_{k=1}^{N} w_{nk} x_{gk} \tag{4}$$

with b_{g0} a good-specific intercept and β_g ($0 \leqslant \beta_g < 1$) a good-specific coefficient, and the w_{nk} *reference weights*, representing the importance attached by consumer n to consumer k's expenditures. They satisfy $w_{nn} = 0$, $w_{nk} \geqslant 0$, and $\Sigma_k w_{nk} = 1$. Intuitively, β_g measures the *conspicuousness* of good g. The higher β_g is, the more one's consumption of good g is influenced by the consumption of others. The expression $\Sigma_{k=1}^{N} w_{nk} x_{gk}$ represents mean expenditures on good g in the *reference group* of consumer n, where the reference group of individual n is defined as the set of individuals k for whom $w_{nk} > 0$.

To allow for differences in household composition, the model is extended slightly in the following simple way. Let f_n be the size of household n, however defined. It is assumed that the household's committed expenditures on good g increase with $\mu_g f_n$, where $\mu_1, \ldots, u_G$ are parameters. This corresponds to 'translating' as defined by Pollak and Wales (1981). Combining preference interdependence with translating leads to the following adaptation of the basic model. Let $\tilde{x}_{gn}$ be defined as

$$\tilde{x}_{gn} \equiv x_{gn} - \mu_g f_n \tag{5}$$

then we replace equation (4) by

$$b_{gn} = b_{g0} + \mu_g f_n + \beta_g \sum_{k=1}^{N} w_{nk} \tilde{x}_{gk} \tag{6}$$

Notice that equation (6) reduces to $b_{gn} = b_{g0} + \mu_g f_n$ in either of two cases: $\beta_g = 0$ or all $\tilde{x}_{gk} = 0$. We may call $b_{g0} + \mu_g f_n$ the *basic needs* of household n, because it represents committed expenditures if the household does not refer to other households at all, or if all other households are just able to satisfy their own basic needs. It is only the excess of other households' expenditures on good g over their basic needs which raises committed expenditures. Incidentally, it would be tempting to call $\tilde{x}_{gn}$ 'discretionary spending' on good g, but we prefer to adhere to the more common definition of discretionary spending as $x_{gn} - b_{gn}$.

The model that we have specified above is closely related to those analysed by Gaertner (1974) and Pollak (1976), although these authors do not take into account demographic variables. Both authors mainly consider dynamic specifications in which the x_{gk} on the right-hand side of equation (4) are lagged one period. The same is done in Alessie and Kapteyn (1991). Gaertner also specifies a relation for γ_g, where an individual's γ_g depends on relative changes in his or her permanent income. Furthermore, he considers various specifications in which the reference weights depend on consumption patterns of individuals. Both Gaertner and Pollak allow the reference weights to vary according to goods and also to be non-zero for $k = n$. Since for our empirical work we only have cross-section data available, a dynamic specification is ruled out. Assuming that the weight w_{nn} an individual gives to his or her own consumption is the same for everyone, it is impossible in a cross-section to distinguish empirically between $w_{nn} = 0$ or $w_{nn} \neq 0$. So, taking $w_{nn} = 0$ ('no habit formation') does not entail loss of generality (although the interpretation of parameter estimates depends on it). Alternatively, it may be argued that $w_{nn} \neq 0$ does not make sense in a static framework.

Also, we do not follow Gaertner's lead to specify a model for γ_g and the w_{nk}, and in contrast to both Gaertner and Pollak the reference weights w_{nk} are assumed identical across commodities up to a constant of proportionality. These are major specifications, inspired by our wish to have a model that can be estimated empirically.

It might be objected to the approach adopted by Gaertner, Pollak and us that there is no clear theoretical reason why the notion of interdependence would be best captured by a linear weighting scheme like equation (4). One could argue that, for example, individuals will refer mainly to others with a higher consumption level. Evidence from social psychology seems to be ambiguous in this respect. For instance, in studies that deal with feelings of equity about remuneration within organizations some find that individuals compare their income to others who are just above them and other studies report that individuals refer mainly to people who are just below them (cf. Von Grumbkov, 1980, and references therein). Altogether, a linear weighting scheme does not seem inconsistent with the available evidence.

Interdependence of preferences has often been invoked to explain altruistic behaviour (e.g. Hochman and Rodgers, 1969, 1973). Also here, one may ask whether it is possible for an individual to increase utility by voluntary transfers of income. Although we do not investigate this issue in detail, we show in footnote 3 that this is not possible under the assumptions made in the next section.

It is worth noticing that the b_{gn} are often interpreted as subsistence levels, so that equation (6) implies that subsistence levels are subject to social influences. In this connection it is of interest to mention some pieces of evidence collected by Smolensky (1965), Ornati (1966), and Mack (published in Miller, 1965, and quoted by Kilpatrick, 1973). In various budget studies, from 1903 to 1960, experts have estimated minimum subsistence levels for the USA. It turns out that the regressions of the log of these subsistence levels on the log of real disposable income per capita in the same year yields elasticities between 0·57 and 0·84. This suggests strongly that, indeed, subsistence levels are subject to social influences. Similarly, responses to the survey question 'What is the smallest amount of money a family of four needs each week to get along in this community?' have shown a trend over time, which is strongly correlated with per capita income (Kilpatrick, 1973; Sawhill, 1988).

3. STOCHASTIC SPECIFICATION

Combining equations (1), (5), and (6) and adding an i.i.d. disturbance term, ε_{gn}, representing all effects on x_{gn} not captured by the systematic part of the model, yields

$$x_{gn} = b_{g0} + \mu_g\left(f_n - \beta_g \sum_{k=1}^{N} w_{nk} f_k\right) + \beta_g \sum_{k=1}^{N} w_{nk} x_{gk} + \gamma_g\left(y_n - \sum_{h=1}^{G} \beta_h \sum_{k=1}^{N} w_{nk} x_{hk}\right)$$
$$+ \gamma_g\left(\sum_{h=1}^{G} \beta_h \mu_h \sum_{k=1}^{N} w_{nk} f_k - \sum_{h=1}^{G} b_{h0} - \tilde{\mu} f_n\right) + \varepsilon_{gn} \qquad (7)$$

where $\tilde{\mu} \equiv \Sigma_{g=1}^{G} \mu_g$. Thus, the model relates expenditures on different goods x_{gn} to total expenditures and family size (y_n and f_n) and expenditures on different goods and family size of others (x_{gk} and f_k) through a linear model with parameters b_{g0}, β_g, γ_g, μ_g, and w_{nk}. The main problem in estimating the model is, of course, created by the large number of reference weights w_{nk}. A related problem is the simultaneity in the system caused by the presence of the x_{gk} on both the left- and the right-hand side.

In earlier work (Van de Stadt, Kapteyn, and Van de Geer, 1985), in a different context, we have adopted the following approach to the estimation of the reference weights w_{nk}. It is intuitively plausible that consumers with a given set of personal characteristics (education, job type, age, etc.) will on average attach a higher weight to expenditures of consumers sharing the same characteristics, than to those of consumers who have different characteristics.[1] This notion can be used to parameterize the weights w_{nk} such that they become a function of the similarity in characteristics between consumers n and k. This function should, of course, contain a much lower number of parameters than $N(N-1)$, the number of linearly independent reference weights. Given such a parameterization, estimation of the newly introduced parameters along with the others becomes feasible, and yields estimates not only of the demand system parameters but also of the reference pattern between groups in society.

Attempts to estimate such reference patterns directly were made by Kapteyn (1977) and Kapteyn, Van Praag, and Van Herwaarden (1978). It leads to very complicated models which are difficult to estimate. The estimates of the parameters describing the pattern of reference weights tend to be unreliable. In this paper we opt for a different, simpler approach: the reference weights are considered to be drawings from a multivariate probability distribution. We do not specify this distribution completely, but make a few assumptions that partly characterize the distribution.

A central concept in our approach is the notion of a *social group*, i.e. a set of people who share certain characteristics such as education, age, job type, etc. The idea is to use the social group to which an individual belongs as a proxy for his or her reference group. To make clear under what circumstances such a procedure is justified and what errors of approximation may be involved, we

[1] It follows from Festinger's theory of social comparison processes (Festinger, 1954) that people will compare primarily to others who are similar, and a large amount of empirical evidence supports this contention to varying degrees. Borrowing from attribution theory, Goethals and Darley (1977) are able to be more specific about how 'similar others' have to be defined. If an individual wants to evaluate a particular ability, for example, he will seek comparison with others who are comparable with respect to attributes related to that ability. Major and Forcey (1985), find that in evaluations of the level of pay received for a job, individuals compare to others who have the same job and sex.

make four explicit assumptions. These four assumptions are listed and discussed in Appendix A. Here we mention only the main implication of the assumptions.

The parameters in equation (7) are estimated by first deriving the reduced form. It turns out that in this reduced form expressions like $\Sigma_k w_{nk}\tilde{y}_k$, where $\tilde{y}_k \equiv y_k - \bar{\mu} f_k$, apear as exogenous variables. The assumptions in Appendix A allow us to approximate these variables as follows (cf. equation (A14)):

$$\sum_k w_{nk}\tilde{y}_k = \kappa\tilde{\eta} + (1-\kappa)\bar{\tilde{y}}_n + \hat{v}_n \tag{8}$$

where $\bar{y}_n$ is the mean of $\tilde{y}_k$ of all families in the social group of individual n, $\tilde{\eta}$ is the mean of all $\tilde{y}_k$ in society, $\hat{v}_n$ is an error term that up to terms of $o_p(1)$[2] is uncorrelated with $\bar{\tilde{y}}_n$ and has mean zero. The interpretation of the parameter κ is that $(1-\kappa)$ is an indicator of the share of the total reference weight that people assign, on average, to others *within* the same social group, whereas κ is the share given, on average, to all people in society, irrespective of whether they are within or outside an individual's social group. So, if $\kappa = 0$, reference groups do not extend beyond one's own social group. If $\kappa = 1$, the social group contains no information whatsoever on one's reference group. In other words, the smaller κ is, the better a proxy one's social group is for one's reference group. Of course, even if $\kappa = 0$ the social group is not a perfect proxy as long as the $\hat{v}_n$ are not identically equal to zero.

Given the approximation (8) the reduced form of (7) takes a simple form, as will appear in the next section.

4. THE REDUCED FORM AND IDENTIFICATION

It is shown in Appendix B that, under the assumptions listed in Appendix A, equation (7) implies the following reduced form:

$$x_{gn} = d_g + \gamma_g y_n + \alpha_g f_n + r_g \bar{y}_n - r_g \bar{\mu} f_n + u_{gn} \tag{9}$$

where $\bar{y}_n$ is mean total consumption and $\bar{f}_n$ is mean family size in the social group of individual n.[3] The reduced-form parameters can be expressed in the structural parameters as follows:

$$r_g \equiv (1-\kappa)\rho_g \tag{10}$$

$$\rho_g \equiv \frac{\beta_g - p}{1-(1-\kappa)\beta_g}\gamma_g \tag{11}$$

$$p \equiv \sum_{g=1}^{G} \frac{\beta_g \gamma_g}{1-(1-\kappa)\beta_g} \Big/ \sum_{g=1}^{G} \frac{\gamma_g}{1-(1-\kappa)\beta_g} \tag{12}$$

$$l_g \equiv \frac{s_g - \varphi\gamma_g}{1-\beta_g} \tag{13}$$

[2] $o_p(1)$ is defined as follows. The random variable x_m is $o_p(1)$ if for any $\varepsilon > 0$, $\lim_{m\to\infty}\Pr(|x_m| > \varepsilon) = 0$. In the present context m refers to the number of individuals in a social group or in society. We shall use $o_p(1)$ for both scalars, vectors and matrices.

[3] One can insert equation (9) into the direct utility function (2) and investigate whether altruistic behaviour by consumer n could improve his utility. It turns out that utility depends positively on terms of the form $\gamma_g y_n - l_g \bar{y}_n$, where l_g is some parameter less than one in absolute value. Altruistic behaviour amounts to a reduction of y_n by Δy_n and an increase of $\bar{y}_n/(N_t - 1)$, where N_t is the number of households in the social group of household n. For N_t sufficiently large, such a redistribution can never increase utility, even if $l_g > 0$.

$$s_g \equiv b_{g0} - \gamma_g \sum_{h=1}^{G} b_{h0} + \kappa\rho_g(\eta - \bar{\mu}\xi) \tag{14}$$

$$\varphi \equiv \sum_{h=1}^{G} \frac{\beta_h s_h}{1-\beta_h} \Big/ \sum_{h=1}^{G} \frac{\gamma_h}{1-\beta_h} \tag{15}$$

$$\alpha_g \equiv \mu_g - \bar{\mu}\gamma_g \tag{16}$$

It is easy to see that α_g, r_g, ρ_g, and d_g add up to zero, when summing over goods. The error term u_{gn} is well behaved in the sense that up to terms of $O(N^{-1})$ it has mean zero and is uncorrelated with the other variables on the right-hand side of equation (9).

Under our assumptions, the reduced-form parameters d_g, γ_g, α_g, r_g, and $\bar{\mu}$ can be estimated consistently from cross-section data (some details follow in Section 5). Knowing, or consistently estimating, the reduced-form parameters does not suffice, however, to determine all structural parameters. This can be seen as follows. Use equation (11) to solve for β_g:

$$\beta_g = \frac{\rho_g + p\gamma_g}{\gamma_g + (1-\kappa)\rho_g} \tag{17}$$

or, with equation (10),

$$1 - (1-\kappa)\beta_g = \frac{\gamma_g(1-(1-\kappa)p)}{\gamma_g + r_g} \tag{18}$$

It follows from the analysis in Appendix B (last paragraph) that, even with κ known, p is unidentified. Since κ is unknown as well, we are lacking two pieces of information for the identification of the β_g. Assuming that $0 \leqslant \kappa < 1$, we are able, however, to infer a *ranking* of β_g's from the reduced-form estimates:

$$\beta_g < \beta_h \Leftrightarrow \frac{r_g}{\gamma_g + r_g} < \frac{r_h}{\gamma_h + r_h} \tag{19}$$

The structural parameters μ_g can be identified from the α_g and $\bar{\mu}$. Notice that without interdependence the μ_g would not be identified, since the α_g sum to zero. Consequently, we would have had only $G-1$ independent pieces of information to identify the G parameters μ_g. It is the presence of $\bar{f}_n$ which makes it possible to identify the sum of the μ_g, $\bar{\mu}$, which provides the extra piece of information required.

The G parameters b_{g0} cannot be identified from equation (13), because the d_g sum to zero. Since the b_{g0} are of no particular interest we do not pay further attention to either the b_{g0} or the d_g.

5. ESTIMATION RESULTS

Model (7) has been estimated using data on 2813 households from the Consumer Expenditure Survey 1981 conducted by the Netherlands Central Bureau of Statistics. As mentioned in Section 2, households have been assigned to social groups with identical characteristics. The characteristics considered are the following:

(1) Educational attainment of head of household (three categories distinguished)
(2) Age of head of household (five categories)
(3) Job type (five categories).

This leads to a maximum of 75 distinct social groups, 56 of which appeared to be represented in the sample.

The variables $\bar{y}_n$ and $\bar{f}_n$ in equation (9) refer to population means in the social group to which individual n belongs. Obvious proxies for $\bar{y}_n$ and $\bar{f}_n$ are the corresponding sample means. Care has been taken, however, for each individual n to base the estimate of $\bar{y}_n$ and $\bar{f}_n$ only on the incomes and family sizes of all *other* sample households in the social group. Of course, replacement of $\bar{y}_n$ and $\bar{f}_n$ by sample means introduces measurement errors, but the variance–covariance matrix of measurement errors in $\bar{y}_n$ and $\bar{f}_n$ corresponding to group t can be estimated unbiasedly by $1/(\tilde{N}_t - 1)$ times the sample covariance matrix of y_n and f_n corresponding to social group t, where $\tilde{N}_t$ is the number of consumers in the sample belonging to social group t.

The model has been estimated by means of the LISREL program (Jöreskog and Sörbom, 1981; Aigner *et al.*, 1984; Bentler, 1983). Under the conditions given in Lemma 4 in Appendix B the LISREL output provides consistent estimates of the reduced-form parameters, and the printed standard errors can serve as asymptotic approximations of the true standard errors of the estimates.

Two sets of estimates of model (7) will be presented, one ignoring the measurement error caused by the use of proxies for $\bar{y}_n$ and $\bar{f}_n$, and one taking into account this measurement error. In the latter case the estimated variance–covariance matrices of the measurement errors per group have been averaged over the groups. (Correlation of measurement error across individuals in the same group has been ignored.) This average error variance–covariance matrix indicates that measurement error accounts for 2·3% of the observed variance of $\bar{y}_n$, for 22·1% of that of $\bar{f}_n$, and for 1·3% of the covariance of $\bar{y}_n$ and $\bar{f}_n$.

Seven expenditure categories are distinguished:

(1) Food
(2) Housing
(3) Clothing
(4) Medical care
(5) Education and entertainment
(6) Transportation
(7) Other expenditures.

The correlation matrix of all variables involved plus their sample means and standard deviations are given in an appendix available upon request.

Because of adding-up restrictions the variance–covariance matrix of the disturbances in model (7) is singular and the parameters satisfy restrictions across equations. As usual, these problems can be accounted for by dropping arbitrarily one of the seven equations (cf. Barten, 1969; Pollak and Wales, 1992). We have chosen to drop the equation for 'other expenditures'.

All money amounts are measured in thousands of guilders per annum. Family size f_n is simply defined as the number of members of household n. Clearly this treatment of family composition is not very sophisticated. The variance–covariance matrix of the reduced-form disturbances of the six maintained equations has been left unrestricted.

The results for different specifications of the model are given in Table I. The χ^2-statistic is an indicator of the extent to which the model is compatible with the data.

Let us first consider the column 'Complete model', which presents the results for the model which takes into account measurement errors in $\bar{y}_n$ and $\bar{f}_n$. According to the χ^2-statistic the model

Table I. Estimation results

Parameter	Complete model	No measurement error	No interdependence
γ_1	0·131 (0·004)	0·131 (0·004)	0·126 (0·003)
γ_2	0·274 (0·006)	0·274 (0·006)	0·287 (0·006)
γ_3	0·081 (0·003)	0·081 (0·003)	0·080 (0·002)
γ_4	0·094 (0·003)	0·094 (0·003)	0·099 (0·002)
γ_5	0·172 (0·015)	0·172 (0·005)	0·171 (0·004)
γ_6	0·238 (0·006)	0·238 (0·006)	0·227 (0·005)
r_1	−0·021 (0·007)	−0·020 (0·007)	
r_2	0·052 (0·013)	0·052 (0·013)	
r_3	−0·003 (0·005)	−0·003 (0·005)	
r_4	0·018 (0·006)	0·018 (0·006)	
r_5	−0·004 (0·009)	−0·004 (0·009)	
r_6	−0·045 (0·012)	−0·045 (0·012)	
μ_1	1·265 (0·154)	1·265 (0·154)	0·729 (0·032)[b]
μ_2	0·828 (0·360)	0·828 (0·361)	−0·317 (0·056)
μ_3	0·492 (0·100)	0·491 (0·100)	0·157 (0·022)
μ_4	0·543 (0·125)	0·543 (0·125)	0·148 (0·024)
μ_5	0·517 (0·212)	0·517 (0·213)	−0·194 (0·041)
μ_6	0·390 (0·276)	0·389 (0·277)	−0·583 (0·051)
$\bar{\mu}$	4·131 (1·223)	4·129 (1·225)	
R_1^2 [a]	0·594	0·594	0·593
R_2^2	0·535	0·535	0·532
R_3^2	0·443	0·443	0·443
R_4^2	0·483	0·483	0·481
R_5^2	0·426	0·426	0·426
R_6^2	0·430	0·430	0·427
χ^2	15·34	15·34	57·76
df	5	5	12

[a] R_g^2 is defined as $1 - \sigma_{ug}^2/\mathrm{var}(x_{gn})$.

[b] The estimates in this column refer to $\alpha_1, \ldots, \alpha_6$.

describes the data well. The estimates of all γ_s are positive and significantly different from zero. Out of the six estimated r_g, four are significantly different from zero.

The column headed 'No measurement error' presents the estimates of the model for the case that the proxies for $\bar{y}_n$ and $\bar{f}_n$ are assumed accurate. This neglect of measurement error does not affect the estimates of the γ_g or the value of the χ^2-statistic up to two decimal places.

The column headed 'No interdependence' presents parameter estimates under the restriction $r_1 = r_2 = \ldots r_6 = 0$. Although the fit of the equations, as gauged by the R^2's, hardly changes and the γ_g and α_g change only marginally, the χ^2-statistic rejects the restrictions decisively. As a final comment on the statistical quality of the results, a χ^2-test of the overidentifying restrictions on the coefficients of f_n and $\bar{f}_n$ does not lead to a rejection.

To start off a discussion of the economic significance of the results, we present information on the structural parameters in Table II. (The last column will be used later.)

Although the β_g are not identified, we can derive their relative ranking from Table II in conjunction with relation (19), assuming that $0 < \kappa < 1$ and all $0 < \beta_g < 1$. We find $\beta_2 > \beta_4 > \beta_5 > \beta_3 > \beta_6 > \beta_1$. Interpreting β_g as a measure of the conspicuousness of good g, we have that the order of conspicuousness is: housing, medical care, education and entertainment,

Table II. Values of structural parameters derived from the reduced-form estimates for the complete model

Expenditure category	$r_g/(\gamma_g + r_g)$	μ_g	γ_g	$\gamma_g + r_g$
1 Food	−0·19	1·27	0·13	0·11
2 Housing	0·16	0·83	0·27	0·33
3 Clothing	−0·04	0·49	0·08	0·08
4 Medical care	0·16	0·54	0·09	0·11
5 Education and entertainment	−0·02	0·52	0·17	0·17
6 Transportation	−0·23	0·39	0·24	0·19

clothing, transportation, food. Except, perhaps, for the relative ranking of medical care and transportation (cars), the ranking seems quite plausible. As to the position of medical care: this name is actually rather a misnomer. Next to medical care it also compromises 'domestic services', a high ranking which seems intuitively plausible. Moreover, there is an artefact at work here, as most households in the sample were compulsorily insured via the sick fund, the contributions to which depend on income and hence, statistically, also on total expenditures in the reference group.

It is of interest to compare predictions of the model for aggregates of all consumers with predictions at the household level. Consider, for instance, an increase of total expenditures by one dollar. At the household level the effect on the expenditures x_{gn} is given by the marginal budget shares γ_g.

It follows from equations (9)–(16) (or more directly by using equations (B4), (B5), and (B7)) that an increase of everyone's total expenditures with one dollar raises each x_{gn} with

$$\frac{\gamma_g}{1 - \beta_g} \bigg/ \sum_h \frac{\gamma_h}{1 - \beta_h} \tag{20}$$

So the extent to which the aggregate consumption of a good responds to changes in total expenditures depends not only on the good's marginal budget share but also on its conspicuousness. One sees that the magnitude of the response for good g is positively related to both its marginal budget share and its conspicuousness.

If $\kappa = 0$ it is easy to show that expression (20) is equal to $\gamma_g + r_g$. Assuming for a moment that $\kappa = 0$, the last two columns of Table II can be used to compare the effect of an increase in total expenditures at the household level and in the aggregate. As one would expect, at the aggregate level effects are larger for conspicuous goods and lower for non-conspicuous goods. Differences can be fairly large. For food, for example, the aggregate effect is about 20% lower than at the household level whereas for housing the aggregate effect is about 20% larger.

In the (standard) model without interdependence individual and aggregate effects are, of course, identical. If one looks at the estimates of the marginal budget shares in the standard model, these appear to be somewhat in between γ_g and $\gamma_g + r_g$ for the complete model. For certain goods the standard model would yield rather misleading predictions of aggregate effects. Taking the complete model as being correct, the standard model would overpredict aggregate effects for food by about 20%, whereas for housing the predictions would be about 10% too low.

Finally we pay attention to the μ's. An easy way to get a feeling for their interpretation is to look at an example. If a family's size increases by one person, the utility function (2) implies that the extra expenditures required on each category to maintain the family's previous utility level are: food, NLG 1200 per annum; housing, NLG 500; clothing, NLG 400; medical care, NLG 400; education and entertainment, NLG 400; transportation, NLG 150. (At the time of the survey, the Dutch guilder, NLG, was approximately US$ 0·50.) These appear to be plausible numbers. But, as always, this equivalence scale interpretation requires arbitrary normalizations. See e.g. Pollak and Wales (1979), Blundell and Lewbel (1991), Pollak and Wales (1979), Kapteyn (1994).

Incidentally, recall from Section 4 that without interdependence the family size coefficients are not identified.

6. CONCLUSIONS

The main purpose of this paper has been to show that preference interdependence can be incorporated into a demand system and to investigate its empirical importance. The results confirm the suspicion that preference interdependence is an important determinant of consumer behaviour, not so much for the extra variance in consumption which can be thus explained nor for the parameter estimates, most of which do not change very much, but certain *conclusions* from the model (e.g. what is the effect of an across-the-board change in total expenditures on the aggregate consumption of various goods?) do change rather substantially if preference interdependence is accounted for. So, to the extent that we want to use a model to predict aggregate responses to changes in exogenous variables, interdependence should not be neglected.

As noted in Section 2, the allocation of total expenditures to a number of expenditure categories can be seen as the second stage in the typical two-stage budgeting process that arises in lifecycle models with intertemporally separable preferences. Here we have restricted ourselves to the second stage, but generally one would surmise that also in the first stage (the determination of the level of total expenditures in any one period) preference interdependence may play a role. Modelling this would probably bring us close to an integration of Duesenberry's relative income hypothesis and Friedman's permanent income hypothesis.

Although spelling out the stochastic assumptions that are required to arrive at a well-behaved reduced form asks for a fair amount of space (Appendix A), and although the derivation of this reduced form is rather tedious (Appendix B), the result is quite simple. Estimation of the model by means of the widely available LISREL computer program is, moreover, straightforward. This suggests that there is really no practical reason to ignore preference interdependence in demand analysis or in other empirical applications of microeconomic theory. Obvious extensions of the analysis in this paper include preference interdependence in labour-supply models (e.g. Kapteyn and Woittiez, 1990), and oligopolistic models of firm behaviour. Lemma 1 of Appendix B provides a rather general framework of the study of interdependence in linear models of interdependent behaviour. Although the restriction to *linear* models may appear severe, it should be noted that a great many of the demand systems used in practical applications are in fact linear in a way that makes Appendix B applicable. Given that we use a cross-section (so that we are in fact estimating Engel curves), the LES is indistinguishable from any other Gorman Polar Form demand system. But also other systems can be extended with preference interdependence in a way that makes Appendix B applicable. In the Almost Ideal Demand System, for instance, one can specify interdependence in terms of budget shares rather than quantities and the same formal structure emerges.

A second extension is to supplement preference interdependence with habit formation. Not only will that probably increase the explanatory power of the model, it will also aid in identifying the structural parameters. This extension requires the availability of panel data.

A third extension has to be in the modelling of reference groups. In this paper we have basically described the distribution of reference weights by means of a single parameter, κ. It should be possible to refine this specification. Ideally, of course, one would like to have a formal theory of how reference groups are formed. To our knowledge no such theory exists at this moment.

APPENDIX A: STOCHASTIC ASSUMPTIONS

Here we introduce and discuss four assumptions that justify the approximation (8) and the reduced form given in Section 3.

If individual n is a member of social group t, $t = 1, \ldots, T$, we denote this as $n \in G_t$, and we denote the size of social group t (i.e. the number of individuals in it) as N_t. As in Section 3, we define $\tilde{y}_k \equiv y_k - \tilde{\mu} f_k$, which is in effect translated total expenditures of household k.

Assumption 1 Within each social group the $\tilde{y}_n$ are random drawings from a probability distribution with mean $\bar{\tilde{y}}_n$, i.e.

$$\tilde{y}_n = \bar{\tilde{y}}_n + \zeta_n \tag{A1}$$

where $\mathrm{E}\zeta_n = 0$; ζ_n is distributed independently from $\bar{y}_n$ and w_{nk} for any n and k.

As a matter of notation, notice that $\bar{\tilde{y}}_n$ is constant within a social group. Sometimes we shall write $\bar{\tilde{y}}_t$ for the value of $\bar{\tilde{y}}_n$ with $n \in G_t$. Define

$$\tilde{\eta} \equiv \frac{1}{N} \sum_t N_t \bar{\tilde{y}}_t \tag{A2}$$

$$\tilde{y}_n^* \equiv \frac{1}{N - N_t} \sum_{s \neq t} N_s \bar{\tilde{y}}_s \qquad n \in G_t \tag{A3}$$

$$p_{ns} \equiv \sum_{k \in G_s} w_{nk} \qquad n \in G_t, s \neq t \tag{A4}$$

$$p_n \equiv \sum_{k \in G_t} w_{nk} \quad n \in G_t \tag{A5}$$

We will refer to $\tilde{\eta}$ as mean translated consumption (mean total expenditures) in society, and to $\tilde{y}_n^*$ as mean translated consumption outside individual n's social group. Obviously, p_{ns} is the total reference weight assigned by individual n to all individuals in social group s, whereas p_n is the total reference weight assigned by this individual to all individuals in his or her own social group.

Assumption 2

$$\sum_{s \neq t} p_{ns}(\bar{\tilde{y}}_s - \tilde{y}_n^*) = \alpha_n(\bar{\tilde{y}}_t - \tilde{\eta}) + o_p(1) \qquad n \in G_t \tag{A6}$$

where α_n is expected to be positive.

The left-hand side of equation (A6) can be interpreted roughly as a covariance between the mean consumption level of a social group and the total reference weight assigned to it. Equation (A6) says that this covariance tends to be positive if individual n belongs to a social group with an above-average level of consumption and that it will be negative if individual n belongs to a social group with a below-average consumption level. The motivation for the assumption is that, generally, one would expect that people will more often tend to assign reference weights to others who are similar (see footnote 1) than to others who are dissimilar. Thus, generally, if someone has a high consumption level he will assign, on average, more weight to others who also have a high consumption level. This induces a correlation between references weights and consumption levels which is positive. On the other hand, someone with a low consumption level will primarily give weights to others with a low consumption level. This induces a negative correlation between reference weights and consumptions. Assumption 2 is a simple way to capture these effects.

As mentioned in Section 3, in the reduced form of the system (7) a variable like $\Sigma_k w_{nk}\tilde{y}_k$ appears. The two assumptions made so far allow us to circumvent the problem of having to specify the reference weights w_{nk}. To see this, first notice that equations (A2) and (A3) imply

$$N\tilde{\eta} = N_t\bar{\bar{y}}_n + (N - N_t)y_n^* \qquad n \in G_t \tag{A7}$$

so that

$$\tilde{y}_n^* = \frac{N}{N - N_t}\tilde{\eta} - \frac{N_t}{N - N_t}\bar{\bar{y}}_n \qquad n \in G_t \tag{A8}$$

Second, define

$$v_n \equiv \sum_k w_{nk}\zeta_k \tag{A9}$$

v_n is a random variable with zero mean and independent of $\bar{\bar{y}}_n, \tilde{y}_n^*$. We have

$$\begin{aligned}
\sum_k w_{nk}\tilde{y}_k &= \sum_k w_{nk}\bar{\bar{y}}_k + v_n \\
&= p_n\bar{\bar{y}}_n + \sum_{s\neq t} p_{ns}\bar{\bar{y}}_s + v_n \\
&= p_n\bar{\bar{y}}_n + \left(\sum_{s\neq t} p_{ns}\right)\tilde{y}_n^* + \sum_{s\neq t} p_{ns}(\bar{\bar{y}}_s - \tilde{y}_n^*) + v_n \\
&= p_n\bar{\bar{y}}_n + (1 - p_n)\tilde{y}_n^* + \alpha_n(\bar{\bar{y}}_n - \tilde{\eta}) + o_p(1) + v_n
\end{aligned} \tag{A10}$$

Using equation (A8) this carries over into

$$\sum_k w_{nk}\tilde{y}_k = (1 - \kappa_n)\bar{\bar{y}}_n + \kappa_n\tilde{\eta} + v_n + o_p(1) \tag{A11}$$

for $n \in G_t$, where

$$\kappa_n \equiv \frac{(1 - p_n)N}{N - N_t} - \alpha_n \tag{A12}$$

To gain some intuition for the meaning of κ_n, let us consider some extreme cases. First, suppose $p_n = 1$, i.e. individual n only assigns weights to others within the same social group. From equation (A6) it is clear that then, generally, $\alpha_n = 0$. Thus $\kappa_n = 0$ and equation (A11) reduces to

$$\sum_k w_{nk}\tilde{y}_k = \bar{\tilde{y}}_n + v_n + o_p(1)$$

Since individual n's reference group is confined within his social group, the social group mean $\bar{\tilde{y}}_n$ is a 'perfect' indicator of the mean consumption in his reference group.

Next, suppose $\kappa_n = 1$. Then equation (A11) reduces to

$$\sum_k w_{nk}\tilde{y}_k = \tilde{\eta} + v_n + o_p(1)$$

Obviously, the social group mean now does not convey any information about the mean consumption in the reference group of individual n.

Basically, equation (A11) reduces the number of unknown parameters from about $N(N-1)$ to about N. A further reduction of the number of unknown parameters is obtained by Assumption 3.

Assumption 3

$$\kappa_n = \kappa + \delta_n \tag{A13}$$

where δ_n is a random variable with mean zero; δ_n and δ_k are independently distributed for $n \neq k$; δ_n is independent of w_{kl} for $k \neq n, l = 1, \ldots, N$.

This assumption mainly serves to further reduce the number of unknown parameters. In particular, it implies a further simplification of equation (A11):

$$\sum w_{nk}\tilde{y}_k = (1-\kappa)\bar{\tilde{y}}_n + \kappa\tilde{\eta} + v_n - \delta_n(\bar{\tilde{y}}_n - \tilde{\eta}) + o_p(1) \tag{A14}$$

Under the above assumptions, $v_n - \delta_n(\bar{\tilde{y}}_n - \tilde{\eta})$ is independent of $\bar{\tilde{y}}_n$. So, rather than having $N(N-1)$ reference weights to deal with we are left with one unknown parameter κ.

To arrive at a reduced form with a well-behaved error term we need one more assumption. Define $w^{(2)}_{nm} \equiv \Sigma_k w_{nk} w_{km}$ and $w^{(l)}_{nm} \equiv \Sigma_k w_{nk} w^{(l-1)}_{km}$ for $l > 2$.

Assumption 4 $\quad \mathrm{E}w^{(l)}_{nm} = O(N^{-1})$, for $l \geqslant 2$

Notice that $w^{(2)}_{nm}$ is the weight assigned by n to m 'via all others'. Assumption 4 therefore states that, on average, the indirect influence of any individual on any other individual will tend to zero if the number of individuals in society tends to infinity.

For the derivation of the reduced form we shall employ the following implication of the assumptions:

$$\sum_k w^{(l)}_{nk}(\bar{\tilde{y}}_k - \tilde{\eta})\delta_k = o_p(1) \qquad \text{for } l \geqslant 2. \tag{A15}$$

The proof of equation (A15) is an application of Chebychev's lemma.

APPENDIX B: DERIVATION OF THE REDUCED FORM

This appendix presents the derivation of the reduced form (9) under the assumption given in Appendix A. We first derive a version of (9) in terms of translated variables, and then adapt the results by including family size. Let

$$
\begin{array}{ll}
\tilde{x} \equiv (\tilde{x}_{11}, \ldots, \tilde{x}_{1N}, \ldots, \tilde{x}_{G1}, \ldots, \tilde{x}_{GN})' & GN \times 1 \\
b \equiv (b_{11}, \ldots, b_{1N}, \ldots, b_{G1}, \ldots, b_{GN})' & GN \times 1 \\
b_0 \equiv (b_{10}, \ldots, b_{G0})' & G \times 1 \\
\gamma \equiv (\gamma_1, \ldots, \gamma_G)' & G \times 1 \\
\tilde{y} \equiv (\tilde{y}_1, \ldots, \tilde{y}_N)' & N \times 1 \\
B \equiv \operatorname{diag}(\beta_1, \ldots, \beta_G)' & G \times G \\
\varepsilon \equiv (\varepsilon_{11}, \ldots, \varepsilon_{1N}, \ldots, \varepsilon_{G1}, \ldots, \varepsilon_{GN})' & GN \times 1 \\
u \equiv (u_{11}, \ldots, u_{1N}, \ldots, u_{G1}, \ldots, u_{GN})' & GN \times 1 \\
W \equiv \begin{bmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & & \vdots \\ w_{N1} & \cdots & w_{NN} \end{bmatrix} & N \times N \\
\bar{\tilde{y}} \equiv (\bar{\tilde{y}}_1, \ldots, \bar{\tilde{y}}_N)' & N \times 1 \\
\rho \equiv (\rho_1, \ldots, \rho_G)' & G \times 1 \\
\beta \equiv (\beta_1, \ldots, \beta_G)' & G \times 1
\end{array}
$$

Let $\imath$ denote a vector of ones, with a subscript that indicates its order. (So, for example, $B\imath_G = \beta$.) Equation (7) can now be written as

$$
\begin{aligned}
\tilde{x} &= b_0 \otimes \imath_N + (B \otimes W)\tilde{x} + \gamma \otimes \{\tilde{y} - (\imath_G' \otimes I_N)(b_0 \otimes \imath_N) + (B \otimes W)\tilde{x}\} + \varepsilon \\
&= c \otimes \imath_N + \{(B - \gamma\beta') \otimes W\}\tilde{x} + \gamma \otimes \tilde{y} + \varepsilon
\end{aligned} \tag{B1}
$$

where c is defined as

$$c \equiv b_0 - \imath_G' b_0 \gamma \tag{B2}$$

Obviously, $\imath_G' c = 0$. Further define

$$A \equiv B - \gamma\beta' \tag{B3}$$

$$\rho \equiv \{I_G - (1 - \kappa)A\}^{-1} A\gamma \tag{B4}$$

$$\psi \equiv (I_G - A)^{-1}(c + \kappa\bar{\eta}A\rho) \tag{B5}$$

$$z \equiv (1 - \kappa)\bar{\tilde{y}} + \kappa\bar{\eta}\imath_N \tag{B6}$$

Furthermore, let

$$\delta \equiv (\delta_1, \ldots, \delta_N)', \quad \bar{\tilde{Y}} \equiv \operatorname{diag}(\bar{\tilde{y}}_1, \ldots, \bar{\tilde{y}}_N), \quad v \equiv (v_1, \ldots, v_N)'$$

Lemma 1 Under assumptions (1)–(3), equation (B1) implies

$$\tilde{x} = \psi \otimes \iota_N + \gamma \otimes \tilde{y} + \rho \otimes z + u \tag{B7}$$

where u satisfies:

$$(I_{GN} - A \otimes W)u = \varepsilon - \rho \otimes (\bar{\tilde{Y}} - \bar{\eta} I_N)\delta + A\gamma \otimes v + o_p(1) \tag{B8}$$

Proof We show that substitution of equation (B7) in (B1) leads to an identity with u satisfying (B8). Equations (B7) and (B1) imply

$$\psi \otimes \iota_N + \gamma \otimes \tilde{y} + \rho \otimes z + u = c \otimes \iota_N + (A \otimes W)(\psi \otimes \iota_N + \gamma \otimes \tilde{y} + \rho \otimes z + u) + \gamma \otimes \tilde{y} + \varepsilon \tag{B9}$$

or, using $W\iota_N = \iota_N$,

$$\psi \otimes \iota_N - c \otimes \iota_N - A\psi \otimes \iota_N + \rho \otimes z - A\gamma \otimes W\tilde{y} - A\rho \otimes Wz + (I_{GN} - A \otimes W)u = \varepsilon \tag{B10}$$

Since, according to equation (B5), $\psi - A\psi = c + \kappa\bar{\eta}A\rho$, the first three terms of the left-hand side of equation (B10) are equal to $\kappa\bar{\eta}A\rho \otimes \iota_N$. So we have

$$(I_{GN} - A \otimes W)u = \varepsilon + A\gamma \otimes W\tilde{y} + A\rho \otimes Wz - \rho \otimes z - \kappa\bar{\eta}A\rho \otimes \iota_N \tag{B11}$$

From equation (A14) and using (B6) we have

$$\begin{aligned} W\tilde{y} &= (1-\kappa)\bar{\tilde{y}} + \kappa\bar{\eta}\iota_N + v - (\bar{\tilde{y}} - \bar{\eta} I_N)\delta + o_p(1) \\ &= z + v - (\bar{\tilde{Y}} - \bar{\eta} I_N)\delta + o_p(1) \end{aligned} \tag{B12}$$

Since from equations (A1) and (A9) $W\bar{\tilde{y}} = W\tilde{y} - v$, we have for Wz (using again equation (B6)):

$$\begin{aligned} Wz &= (1-\kappa)W\bar{\tilde{y}} + \kappa\bar{\eta}\iota_N \\ &= (1-\kappa)\{z - (\bar{\tilde{Y}} - \bar{\eta} I_N)\delta + o_p(1)\} + \kappa\bar{\eta}\iota_N \end{aligned} \tag{B13}$$

Collecting terms, we find for equation (B11):

$$\begin{aligned} (I_{GN} - A \otimes W)u = {} & A\gamma \otimes z + (1-\kappa)A\rho \otimes z - \rho \otimes z + \varepsilon + A\gamma \otimes v \\ & - (A\gamma + (1-\kappa)A\rho) \otimes (\bar{\tilde{y}} - \bar{\eta} I_N)\delta + o_p(1) \end{aligned} \tag{B14}$$

It follows immediately from equation (B4) that $A\gamma + (1-\kappa)A\rho = \rho$. Hence equation (B14) simplifies to (B8). □

Lemma 2

$$(I_{GN} - A \otimes W)^{-1} = \sum_{j=0}^{\infty} A^j \otimes W^j \tag{B15}$$

Proof For any integer $l > 1$:

$$\left(\sum_{j=0}^{l-1} A^j \otimes W^j\right)(I_{GN} - A \otimes W) = I_{GN} - A^l \otimes W^l \tag{B16}$$

So to prove equation (B15) it is sufficient to prove that $A^l \otimes W^l$ converges to zero if l tends to infinity. Since W is a Markov matrix, W^l is a Markov matrix as well. Hence the elements of W^l are bounded (they have values between zero and one). It is therefore sufficient to prove that $A^l \to 0$ for l to infinity. We show this by proving that the eigenvalues of A are all within the unit circle (Oldenburger, 1940).

First, assume that all β_g are different and strictly positive. Then the eigenvalues of A follow from the determinantal equation

$$\begin{aligned} |A - \lambda I_G| &= |B - \lambda I_G - \gamma\beta'| \\ &= |\beta - \lambda I_G|\{1 - \beta'(B - \lambda I_G)^{-1}\gamma\} \\ &= 0 \end{aligned} \tag{B17}$$

The expression between braces equals

$$\iota_G'(B - \lambda I_G)(B - \lambda I_G)^{-1}\gamma - \beta'(B - \lambda I_G)^{-1}\gamma = -\lambda\iota_G'(B - \lambda I_G)^{-1}\gamma \tag{B18}$$

so $\lambda = 0$ or $\iota_G'(B - \lambda I_G)^{-1}\gamma = 0$. In scalar notation, the latter expression reads

$$\sum_{g=1}^{G} \frac{\gamma_g}{\beta_g - \lambda} = 0 \tag{B19}$$

Each of the terms under the summation sign is an orthogonal hyperbola in λ with $\lambda = \beta_g$ as its vertical asymptote. So equation (B19) has a solution between each two successive β_g, giving the remaining $G - 1$ roots of equation (B17). So all roots are non-negative and smaller than the largest β_g, which, by assumption, is less than 1.

This still holds when not all β_g are different or strictly positive. This follows directly from the continuity of eigenvalues of a matrix as a function of its elements. □

Lemma 3 Under assumptions (1)–(4) and ignoring terms of order $o_p(1)$, the vector u satisfying equation (B10) has mean zero and

$$\mathrm{E}u\bar{y}' = A\gamma \otimes \mathrm{E}W + O(N^{-1}) \tag{B20}$$

Proof Use Lemma 2 to rewrite equation (B8) as

$$\begin{aligned} u = {} & (I_{GN} - A \otimes W)^{-1}\{\varepsilon + A\gamma \otimes v\} \\ & - (I_G + A)\rho \otimes W(\bar{\bar{Y}} - \bar{\eta}I_N)\delta \\ & - \sum_{j=2}^{\infty} A^j\rho \otimes W^j(\bar{\bar{Y}} - \bar{\eta}I_N)\delta + o_p(1) \end{aligned} \tag{B21}$$

Now, equation (A15) implies

$$\sum_{j=2}^{\infty} A^j \rho \otimes W^j(\tilde{\bar{y}} - \tilde{\eta} I_N)\delta = \sum_{j=2}^{\infty} A^j \rho \otimes o_p(1)$$
$$= A^2(I_G - A)^{-1}\rho \otimes o_p(1)$$
$$= o_p(1) \tag{B22}$$

So we have for u:

$$u = (I_{GN} - A \otimes W)^{-1}(\varepsilon + A\gamma \otimes v) - (I_G + A)\rho \otimes W(\tilde{\bar{Y}} - \tilde{\eta} I_n)\delta + o_p(1) \tag{B23}$$

The first term on the right-hand side involves ε and $v \equiv W\zeta$ where $\zeta \equiv (\zeta_1, \ldots, \zeta_N)'$. Since both ε and ζ are independent of W and have expectation zero, this first term has expectation zero as well. In the second term the random variables are W and δ. A typical element of $W(\tilde{\bar{Y}} - \tilde{\eta} I_N)\delta$ is $\Sigma_{k \neq n}(\tilde{\bar{y}}_k - \tilde{\eta}) w_{nk}\delta_k$. Since δ_k has mean zero and is independent of w_{nk} for $k \neq w$, this element has mean zero. Consequently the second term has mean zero. Neglecting the $o_p(1)$ term, we conclude that u has mean zero.

To prove the second part of the lemma, we first observe that δ and ε are independent of y. So we only have to consider

$$\mathrm{E}(I_{GN} - A \otimes W)^{-1}(A\gamma \otimes v)\tilde{y}' = \mathrm{E}(I_{GN} - A \otimes W)^{-1}(A\gamma \otimes W\tilde{y}\tilde{y})'$$
$$= \mathrm{E}(I_{GN} - A \otimes W)^{-1}(A\gamma \otimes W\mathrm{E}\tilde{y}\tilde{y}')$$
$$= \sigma_y^2 \mathrm{E}(I_{GN} - A \otimes W)^{-1}(A\gamma \otimes W) \tag{B24}$$

where the second equality sign is based on the independence of W and ζ. Next we write

$$\mathrm{E}(I_{GN} - A \otimes W)^{-1}(A\gamma \otimes W) = \mathrm{E}\sum_{j=1}^{\infty} A^j\gamma \otimes W^j$$
$$= \sum_{j=1}^{\infty} A^j\gamma \otimes \mathrm{E}W^j$$
$$= A\gamma \otimes \mathrm{E}W + \sum_{j=2}^{\infty} A^j\gamma \otimes \mathrm{E}W^j$$
$$= A\gamma \otimes \mathrm{E}W + \sum_{j=2}^{\infty} A^j\gamma \otimes O(N^{-1})$$
$$= A\gamma \otimes \mathrm{E}W + A^2(I - A)^{-1}\gamma \otimes O(N^{-1})$$
$$= A\gamma \otimes \mathrm{E}W + O(N^{-1}) \tag{B25}$$

where the third equality follows from Assumption 4. □

Note that the diagonal elements of W are identically equal to zero. As a result, an element of u corresponding to a certain observation is uncorrelated with the element of $\tilde{y}$ corresponding to that same observation. Of course, any element of u does correlate with elements of $\tilde{y}$ corresponding to *different* observations, but that does not affect the asymptotic distribution of the ML-estimator. This statement is made somewhat more precise in Lemma 4.

Let us define the 'conventional' ML-estimator of the reduced-form parameters ψ, γ, and ρ in equation (B7) as the estimator that maximizes the likelihood of the observations under the assumption that u follows a normal distribution (with mean zero) with a variance–covariance matrix of the form $\Sigma \otimes I_N$, where Σ is unrestricted. This estimator provides us with consistent estimates of ψ, γ, and ρ under assumptions (1)–(4), but in order to use the corresponding conventional standard errors an extra assumption is needed. This is summarized, somewhat informally, in Lemma 4.

Lemma 4 Under assumptions (1)–(4), the conventional ML-estimator of the reduced-form parameters is consistent. If we strengthen Assumption 4 to

$$\mathrm{E}w_{nm} = O(N^{-1})$$

then the conventional standard errors are consistent estimates of the true standard errors.

Proof To prove the first part, ignore for a moment the overidentifying restriction implicit in the definitions of $\tilde{x}$ and $\tilde{y}$ (both depend on the μ's). Then equation (B7) is simply a system of seemingly unrelated regressions where the same explanatory variables appear in each equation. Consequently, the conventional ML-estimator is identical to the OLS-estimator applied equation by equation. Since the diagonal elements of W are identically equal to zero, it follows from Lemma 3 that the elements of u are uncorrelated with the explanatory variables corresponding to the same observation. It follows immediately the OLS-estimator is consistent. Now, taking into account the overidentifying restrictions does not impair consistency.

Concerning the second part of the lemma, we observe that the strengthened version of Assumption 4 in conjunction with Lemma 3 implies that we can neglect the correlation between u and $\tilde{y}$. Furthermore, considering equation (B21) it is clear that the only source of correlation of elements of u across observations arises from terms involving W. By assumption these terms can be neglected. As a result u has the variance–covariance matrix assumed by the conventional ML-estimator and its standard errors are consistent estimates of the true standard errors of the parameter estimates. □

To conclude the derivation of the reduced form, we rewrite equation (B7) in terms of non-translated variables. Define:

$$d \equiv \psi + \tilde{\eta}\kappa\rho \tag{B26}$$

$$\alpha \equiv \mu - \tilde{\mu}\gamma \tag{B27}$$

$$r \equiv (1 - \kappa)\rho \tag{B28}$$

Now we have

Lemma 5 Under assumptions (1)–(4), equation (B7) implies

$$x = d \otimes \iota_N + \alpha \otimes f + \gamma \otimes y + r \otimes \tilde{y} - r\tilde{\mu} \otimes \bar{f} + u \tag{B29}$$

Up to terms of $O(N^{-1})$ the error term u has mean zero and u_{gn} is uncorrelated with y_n and f_n. Lemma 4 applies.

The error term u satisfies an expression similar to equation (B10) and the properties of u follow from arguments similar to Lemma 3. It is also a matter of analogy to prove that Lemma 4 applies, except for one slight complication. In equation (B29) there are overidentifying restrictions on the reduced-form parameters so that ML is no longer identical to OLS equation by equation. Since imposition of correct restrictions does not impair consistency, the consistency of the ML-estimator still follows from the consistency of the OLS-estimator.

Finally we substantiate the remark following equation (18) in Section 4, by using equation (B4) to express β as a function of γ, ρ, and κ. First rewrite equation (B4) as

$$A\gamma + (1-\kappa)A\rho = \rho \tag{B30}$$

or

$$(I_G - \gamma\iota_G')B(\gamma + (1-\kappa)\rho) = \rho \tag{B31}$$

Let Δ be the diagonal matrix with typical diagonal element $\gamma_g + (1-\kappa)\rho_g$. Then equation (B31) is eauivalent to

$$(I_G - \gamma\iota_G')\Delta\beta = \rho \tag{B32}$$

As $I_G - \gamma\iota_G'$ has rank $G-1$, $\Delta\beta$ cannot uniquely be inferred from equation (B32). Using the algebra of singular linear systems (e.g. Searle, 1971), the general solution of equation (B32) is, for arbitrary p,

$$\Delta\beta = (I_G - \gamma\iota_G')\rho + p\gamma = \rho + p\gamma \tag{B33}$$

or

$$\beta = \Delta^{-1}(\rho + p\gamma) \tag{B34}$$

This is equivalent to equation (17).

ACKNOWLEDGMENTS

Financial support by the Netherlands Organization for Scientific Research is gratefully acknowledged. We thank Ton Barten, Wouter Keller, Peter Kooreman, Rob Alessie, and two anonymous referees for valuable comments and Erik van Elderen for research assistance.

REFERENCES

Aigner, D. J., C. Hsiao, A. Kapteyn and T. J. Wansbeek (1984), 'Latent variable models in econometrics', in Z. Griliches and M. D. Intriligator (eds), *Handbook of Econometrics*, North-Holland, Amsterdam, vol. 2, 1331–93.

Alessie, R. J. M. and A. Kapteyn (1989), 'Habit formation, interdependent preferences and demographic effects in the Almost Ideal Demand System', *The Economic Journal*, **101**, 404–19.

Alessie, R. J. M. and A. Kapteyn (1989), 'Consumption, savings and demography', in A. Wenig and K. F. Zimmerman (eds), *Demographic Change and Economic Development*, Springer, Berlin, 272–305.
Alessie, R. J. M. and A. Kapteyn (1991), 'Habit formation, interdependent preferences and demographic effects in the almost ideal demand system', *The Economic Journal*, **101**, 404–19.
Barten, A. P. (1969), 'Maximum likelihood estimation of a complete system of demand equations', *European Economic Review*, **1**, 7–73.
Bentler, P. M. (1983), 'Simultaneous equation-systems as moment structure models; with an introduction to latent variable models', *Journal of Econometrics*, **22**, 13–42.
Blanciforti, L. and R. Green (1983), 'An almost ideal demand system incorporating habits: An analysis of expenditures on food and aggregate commodity groups', *The Review of Economics and Statistics*, **65**, 511–15.
Blomquist, N. S. (1993), 'Interdependent behavior and the effect of taxes', *Journal of Public Economics*, **51**, 211–18.
Blundell, R. A. and A. Lewbel (1991), 'The information content of equivalence scales', *Journal of Econometrics*, **50**, 49–68.
Blundell, R. A. and I. Walker (1986), 'A life cycle consistent empirical model of family labour supply using cross-section data', *Review of Economic Studies*, **53**, 539–58.
Darrough, M., R. A. Pollak and T. J. Wales (1983), 'Dynamic and stochastic structure: An analysis of three time series of household budget studies', *The Review of Economics and Statistics*, **65**, 274–81.
Duesenberry, J. S. (1949), *Income, Saving and the Theory of Consumer Behavior*, Harvard University Press, Cambridge, MA.
Festinger, L. (1954), 'A theory of social comparison processes', *Human Relations*, **7**, 117–40.
Frank, R. H. (1984), 'Interdependent preferences and the competitive wage structure', *Rand Journal of Economics*, **15**, 510–20.
Frank, R. H. (1985), *Choosing the Right Pond*, Oxford University Press, Oxford.
Frank, R. H. (1989), 'Frames of reference and the quality of life', *American Economic Review*, **79**, 80–5.
Gaertner, W. (1974), 'A dynamic model of interdependent consumer behavior', *Zeitschrift für Nationalökonomie*, **34**, 327–44.
Goethals, G. R. and J. M. Darley (1977), 'Social comparison theory. An attributional approach', in J. M. Suls and R. L. Miller (eds), *Social Comparison Processes*, Hemisphere, New York.
Hochman, H. M. and J. D. Rodgers (1969), 'Pareto optimal redistribution', *The American Economic Review*, **59**, 542–57.
Hochman, H. M. and J. D. Rogers (1973), 'Utility interdependence and income transfers through charity', in K. E. Boulding and M. Phaff (eds), *Transfers in an Urbanized Economy*, Wadsworth, Belmont, CA.
Houthakker, H. S. and L. D. Taylor (1970), *Consumer Demand in the United States 1929–1970, Analysis and Projections*, Harvard University Press, Cambridge, MA.
Jöreskog, K. G. and D. Sörbom (1981), 'LISREL VI — Analysis of linear structural relationships by maximum likelihood, instrumental variables and least squares methods', Department of Statistics, University of Uppsala.
Kapteyn, A. (1977), *A Theory of Preference Information*, PhD thesis, Leiden University.
Kapteyn, A. (1994), 'The measurement of household cost functions; revealed preference versus subjective measures', *Journal of Population Economics*, **7**, 333–50.
Kapteyn, A. and I. Woittiez (1990), 'Preference interdependence and habit formation in family labour supply', in J. P. Florens *et al.* (eds), *Microeconometrics: Surveys and applications*, Tieto, Clevedon.
Kapteyn, A., B. M. S. van Praag and F. G. van Herwaarden (1978), 'Individual welfare functions and social reference spaces', *Economics Letters*, **1**, 173–78.
Kapteyn, A., T. J. Wansbeek and J. Buyze (1980), 'The dynamics of preference formation', *Journal of Economic Behavior and Organization*, **1**, 123–57.
Kilpatrick, R. W. (1973), 'The income elasticity of the poverty line', *The Review of Economics and Statistics*, **55**, 327–32.
Leibenstein, H. (1950), 'Bandwagon, snob, and Veblen effects in the theory of consumer's demand', *Quarterly Journal of Economics*, **64**, 183–207.
Major, B. and B. Forcey (1985), 'Social comparisons and pay evaluations: preferences for same-sex and same-job wage comparisons', *Journal of Experimental Social Psychology*, **21**, 393–405.

Manser, M. E. (1976), 'Elasticities of demand for food: An analysis using non-additive utility functions allowing for habit formation', *Southern Economic Journal*, **43**, 879–91.
Manski, C. F. (1993), 'Identification of endogenous social effects: the reflection problem', *The Review of Economic Studies*, **60**, 531–42.
Miller, H. P. (1965), 'Changes in the number and composition of the poor', in: M. S. Gordon (ed.), *Poverty in America*, Chandler, San Francisco, 81–101.
Muellbauer, J. (1988), 'Habits, rationality, and myopia in the life-cycle consumption function', *Annales d'Economie et de Statistique*, **9**, 47–70.
Oldenburger, R. (1940), 'Infinite powers of matrices and characteristic roots', *Duke Mathematical Journal*, **6**, 357–61.
Ornati, O. (1966), *Poverty amid Affluence*, Twentieth Century Fund, New York.
Pashardes, P. (1986), 'Myopic and forward looking behaviour in a dynamic demand system', *International Economic Review*, **27**, 387–97.
Phlips, L. (1972), 'A dynamic version of the linear expenditure model', *The Review of Economics and Statistics*, **54**, 450–8.
Phlips, L. (1974), *Applied Consumption Analysis*, North-Holland, Amsterdam.
Phlips, L. and F. Spinnewijn (1982), 'Rationality versus myopia in dynamic demand systems', *Advances in Econometrics*, **1**, 3–33.
Pollak, R. A. (1976), 'Interdependent preferences', *American Economic Review*, **66**, 309–20.
Pollak, R. A. and T. J. Wales (1969), 'Estimation of the linear expenditure system', *Econometrica*, **37**, 611–28.
Pollak, R. A. and T. J. Wales (1979), 'Welfare comparisons and equivalence scales', *American Economic Review*, **69**, 216–21.
Pollak, R. A. and T. J. Wales (1981), 'Demographic variables in demand analysis', *Econometrica*, **49**, 1533–52.
Pollak, R. A. and T. J. Wales (1992), *Demand System Specification and Estimation*, Oxford University Press, Oxford.
Sawhill, I. V. (1988), 'Poverty in the U.S.: why is it so persistent?' *Journal of Economic Literature*, **26**, 1073–119.
Searle, S. R. (1971), *Linear Models*, Wiley, New York.
Smolensky, E. (1965), 'The past and present poor', in *The Concept of Poverty*, US Chamber of Commerce, New York, 35–67.
Spinnewijn, F. (1981), 'Rational habit formation', *European Economic Review*, **15**, 91–109.
Van de Stadt, H., A. Kapteyn and S. van de Geer (1985), 'The relativity of utility: evidence from panel data', *The Review of Economics and Statistics*, **67**, 179–87.
Von Grumbkov, J. (1980), *Sociale vergelijking van salarissen*, PhD thesis, University of Groningen.
Winder, C. C. A. and F. C. Palm (1989), 'Intertemporal consumer behaviour under structural changes in income', *Econometric Reviews*, **8**, 1, 1–88.